GSFC Earth Sciences Distributed Active Archive Center (GES DAAC)

Gary Alcott, Steve Kempler, Greg Leptoukh, Chris Lynnes, Long Pham, Peter Smith, Bruce Vollmer, Gail Wade

Presentation Purpose

To Discuss:

- What is the GES DAAC?
- What does the GES DAAC do?

Note:

The GES DAAC includes activities directly funded by ESDIS Project (e.g, EOS missions)

The GES Data and Information Services Center (DISC) includes DAAC activities, plus those not directly funded by the ESDIS Project (e.g., ACCESS proposals)

This UWG will address DAAC activities.

GES DAAC Mission

The GES DAACs mission is to maximize NASA's investment benefit by providing data and services that enable people to fully realize the scientific and educational potential of global climate data.

In Short...

The GES DAACs mission is to:

ENABLE EARTH SCIENCE

Summary of Expertise

- **Software engineering** - *In-house expertise and experience that understands the best advanced technologies to further mature data management system usability and efficiency*
- **Science data management expertise** – *Doctorate-level scientists in interdisciplinary Earth sciences who collaborate with researchers to develop sophisticated Web-based tools to facilitate comprehensive information management, access, analysis and visualization*
- **Mission Support** - *In-house expertise who understands the requirements for costing and sizing information management systems for new or existing missions*
- **Operational active archive and distribution system with complete user services** - *8 X5 (24 X 7 on call) staff that understands the importance of, and ensures, continuous data ingest, processing, archive and distribution*

Missions Supported

Atmospheric Composition

- Total Ozone Mapping Spectrometer (TOMS)
- Upper Atmosphere Research Satellite (UARS)
- Solar Radiation and Climate Experiment (SORCE)
- Aura: Ozone Monitoring Instrument (OMI), High Resolution Dynamics Infrared Sounder (HIRDLS), Microwave Limb Sounder (MLS)
- *Soon*: GLORY: Aerosol Polarimetry Sensor (APS), Total Irradiance Monitor (TIM), Cloud Camera

Atmospheric Dynamics

- TIROS Operational Vertical Sounder (TOVS)
- Aqua Atmospheric Infrared Sounder (AIRS)

Modeling

- Data Assimilation Office (DAO), Global Modeling Assimilation Office (GMAO)
- Land Data Assimilation System (LDAS)

Precipitation

- Tropical Rainfall Measuring Mission (TRMM)
- Hydrology Data Collections

Previously missions supported

Terra and Aqua: Moderate Resolution Imaging Spectroradiometer (MODIS)

Coastal Zone Color Scanner (CZCS); Sea-viewing Wide-field-of-view Sensor (SeaWIFS)

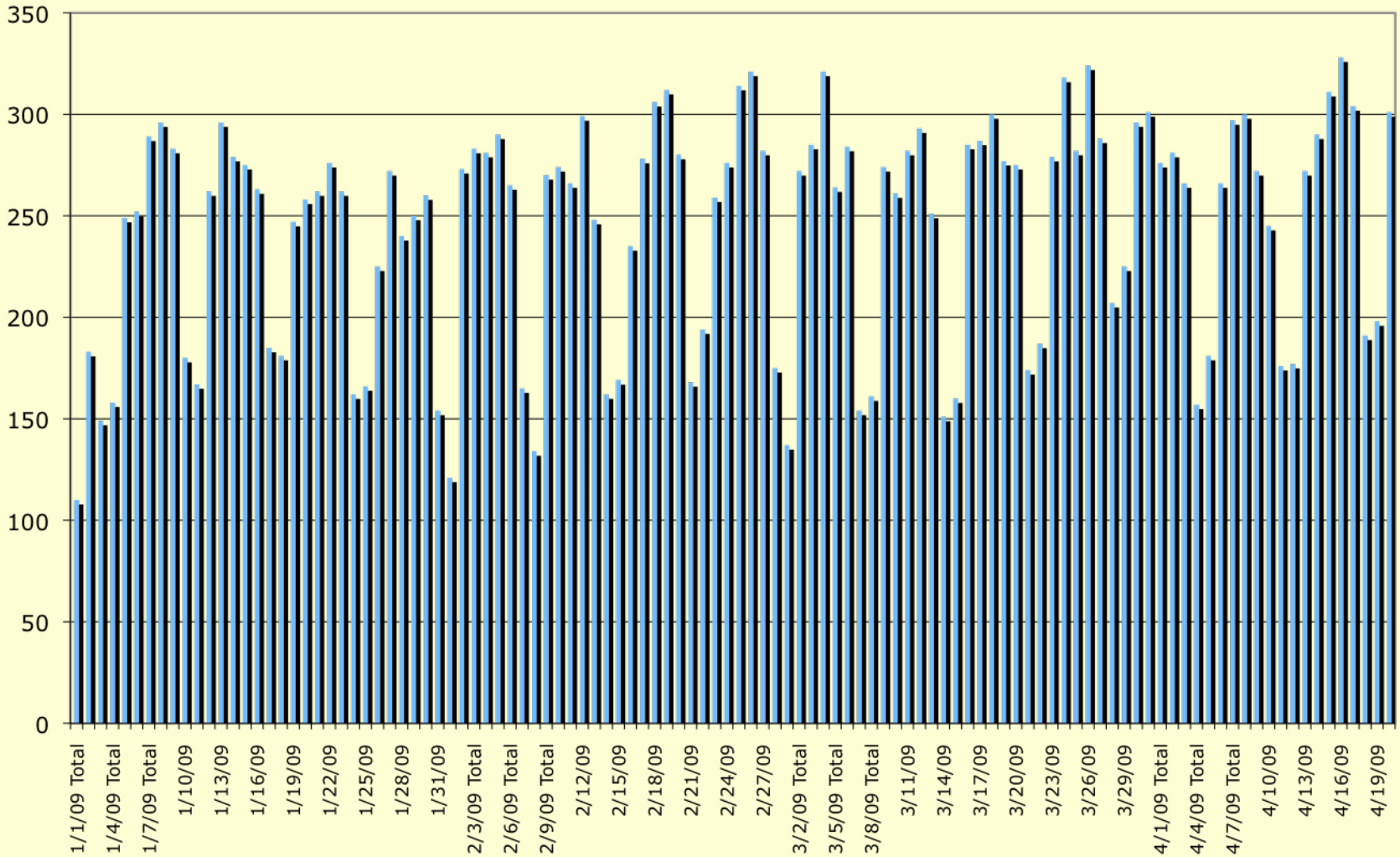
Then .. And Now

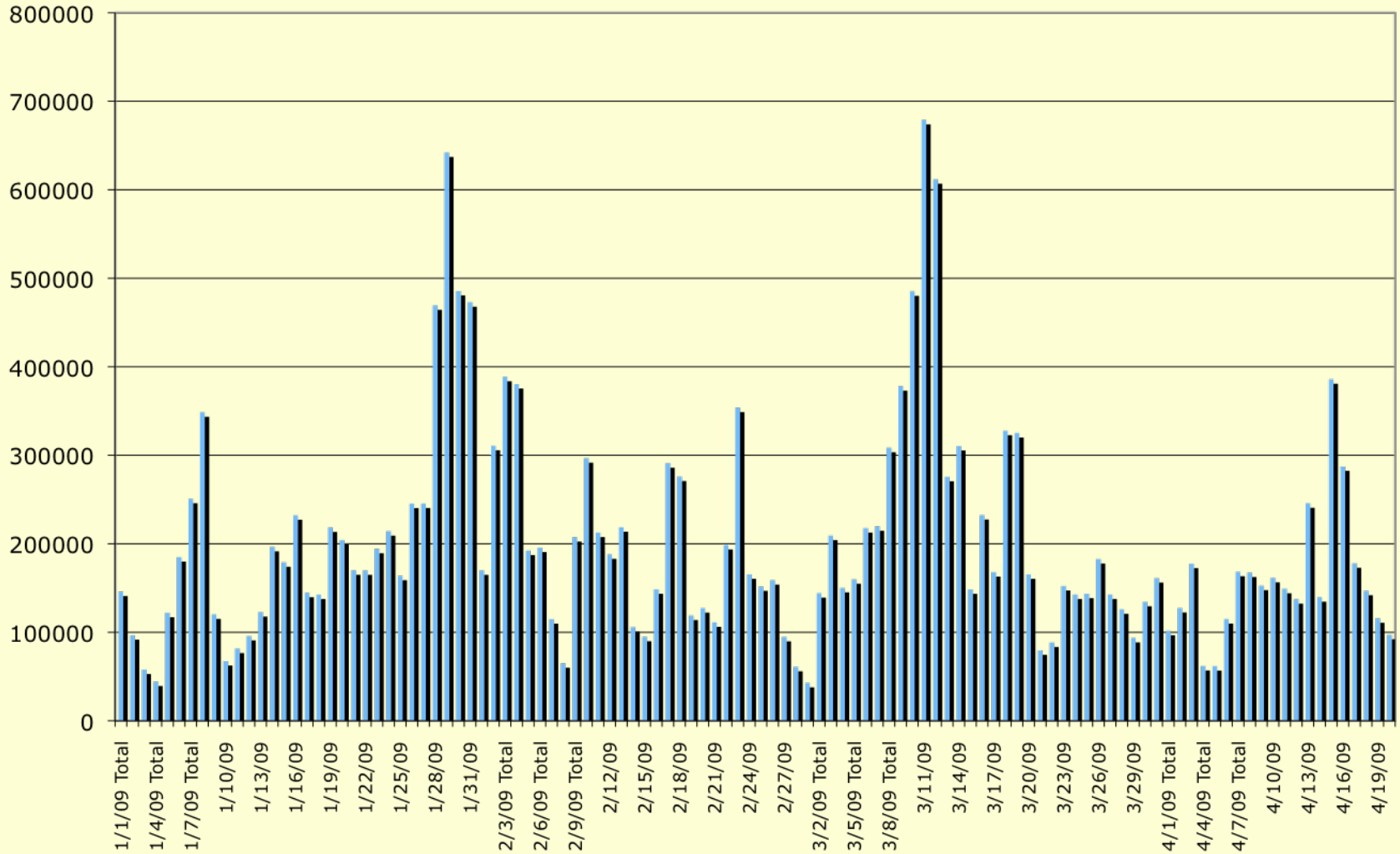
(Does not include non-EOSDIS funding)

	2000	2004	2009
EOSDIS funding ($K)	5,553	13,301	7,435
EOSDIS FTEs	CS = 10 Contractor = 51	CS = 9 Contractor = 103	CS = 8 Contractor = 49
Vol (TB)	100	1700	250 (single copy), 400 (total)
Instruments	CZCS, TOMS, TOVS, DAO, AVHRR, UARS, SeaWIFS, TRMM, Terra (MODIS)	CZCS, TOMS, TOVS, DAO, UARS, SeaWIFS, TRMM, Terra (MODIS), **Aqua (MODIS,AIRS), SORCE, Aura (OMI, HIRDLS, MLS)**	TOMS, TOVS, GMAO, UARS, TRMM, Aqua (AIRS), SORCE, Aura (OMI, HIRDLS, MLS), **LDAS, A-Train subsetted datasets, Hurricane data (TRMM, QuikScat)**
Main Services	Subsetting, Data Ordering, S4P	Subsetting, Subsampling, Data Ordering, Data Mining, S4PM, TOVAS	Subsetting, Subsampling, Data Ordering, Data Mining, S4PM, S4PA, Giovanni, Mirador, Data Fusion, Multi-Sensor Data Co-Registration
Archive Type	Near-line tape, Disk	Near-line tape, Disk	Disk
Systems	Version 0, Version 1, ECS	S4PA, ECS	S4PA

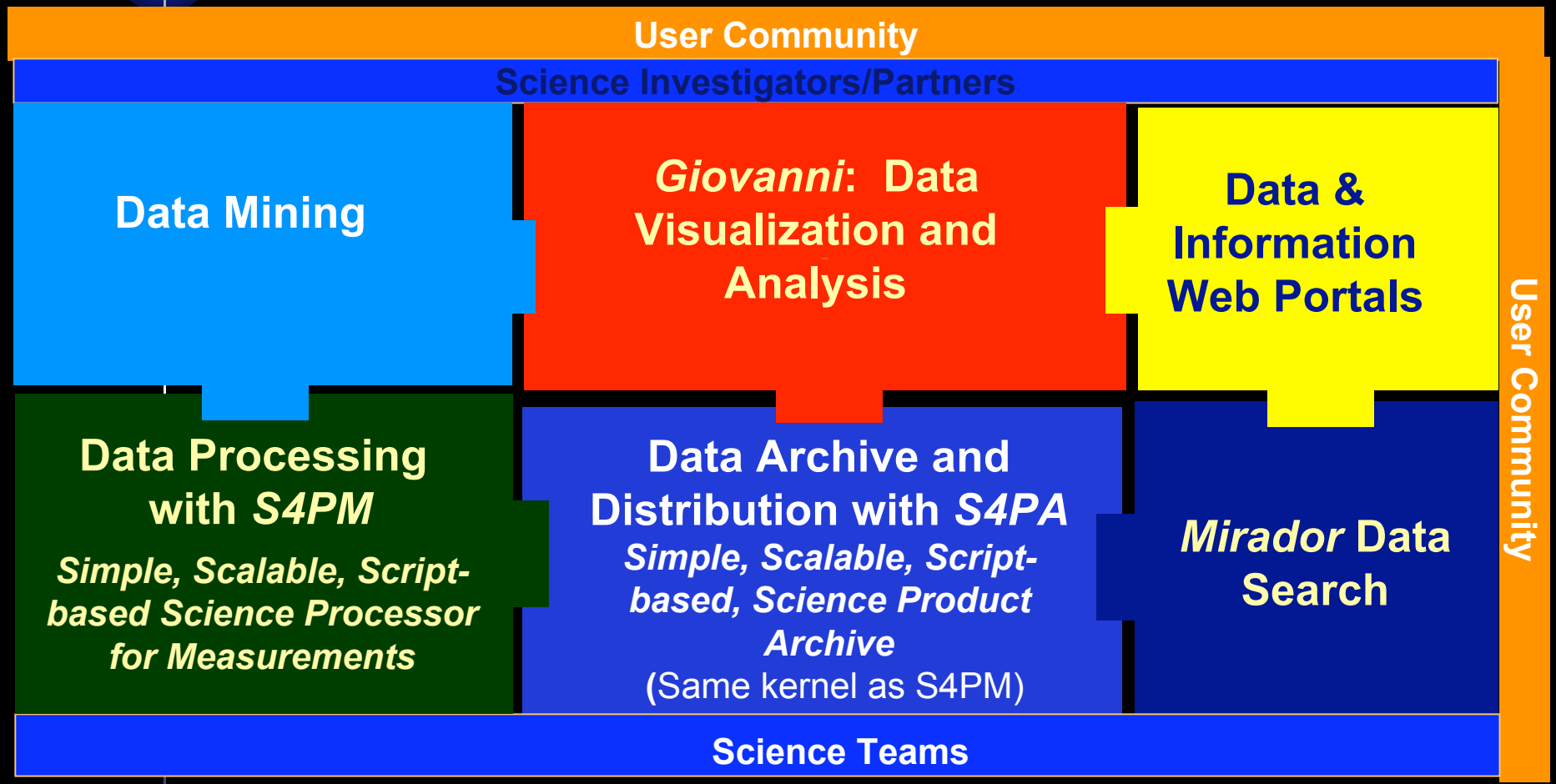
EOSDIS Evolution at the GES DAAC: 2006-2007

Pre-Evolution System Characteristics - ECS based	Evolved System Characteristics - S4PA based
Generalized interface(s)	Discipline-specific interfaces in addition to generalized interface(s)
Tape archive - All products archived - Order data for delivery	Disk archive - Some products processed on demand (virtual products) - Download data automatically upon choosing
Search and order tools	Tools to find, explore, and analyze data
Distribute standard products	Distribute lower volume tailored products
System changes require long lead time	System changes implemented quickly according to priority within given budget cap
Steward data	Steward data


Distibution Public Users (Distinct IPs per day)
350
300
250
200
150
100
50
0
1/1/09 Total
1/4/09 Total
1/7/09 Total
1/10/09
1/13/09
1/16/09
1/19/09
1/22/09
1/25/09
1/28/09
1/31/09
2/3/09 Total
2/6/09 Total
2/9/09 Total
2/12/09
2/15/09
2/18/09
2/21/09
2/24/09
2/27/09
3/2/09 Total
3/5/09 Total
3/8/09 Total
3/11/09
3/14/09
3/17/09
3/20/09
3/23/09
3/26/09
3/29/09
4/1/09 Total
4/4/09 Total
4/7/09 Total
4/10/09
4/13/09
4/16/09
4/19/09


Distribution Granules via FTP (per day)
800000
700000
600000
500000
400000
300000
200000
100000
0
1/1/09 Total
1/4/09 Total
1/7/09 Total
1/10/09
1/13/09
1/16/09
1/19/09
1/22/09
1/25/09
1/28/09
1/31/09
2/3/09 Total
2/6/09 Total
2/9/09 Total
2/12/09
2/15/09
2/18/09
2/21/09
2/24/09
2/27/09
3/2/09 Total
3/5/09 Total
3/8/09 Total
3/11/09
3/14/09
3/17/09
3/20/09
3/23/09
3/26/09
3/29/09
4/1/09 Total
4/4/09 Total
4/7/09 Total
4/10/09
4/13/09
4/16/09
4/19/09

GES DAAC Core Components



Data Archive and Distribution with *S4PA*

Simple, Scalable, Script-based, Science Product Archive

- **Radically simplified architecture for archive and distribution**
- **Features**
  - **Public and restricted-access**
  - **Subscriptions**
  - **Automated data integrity checking**

Interfaces to:

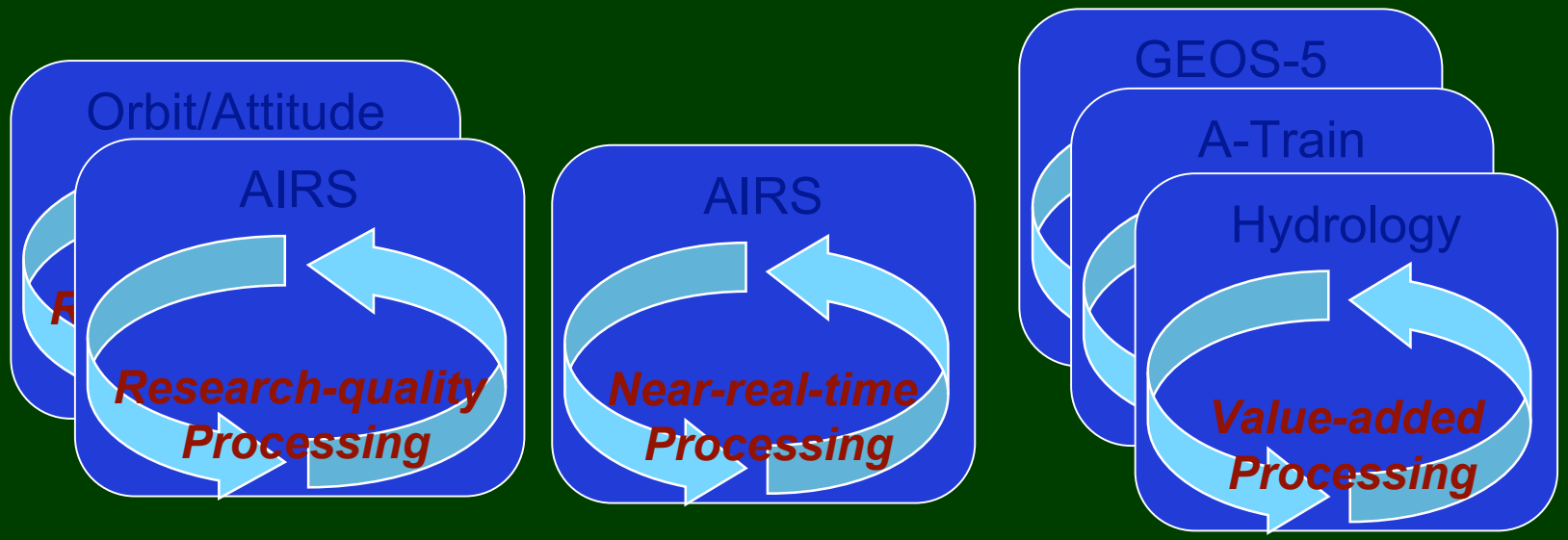
- **S4PM processing system**
- **Science Investigator Processing Systems (SIPS)**
- **EOS Data Operations (EDOS)**
- **Mirador search tool**
- **Giovanni**
- **Global Change Master Directory**
- **EOS Clearinghouse**

Data Processing with *S4PM*

Simple, Scalable, Script-based Science Processor for Measurements

- In-house developed open-source software
- Runs all data processing at the GES DISC since 2002
- Near-real-time processing: AIRS, MLS/Aura (in progress)
- Reused by LaRC for CALIPSO, FlashFlux, MISR, EDC for ASTER On-Demand

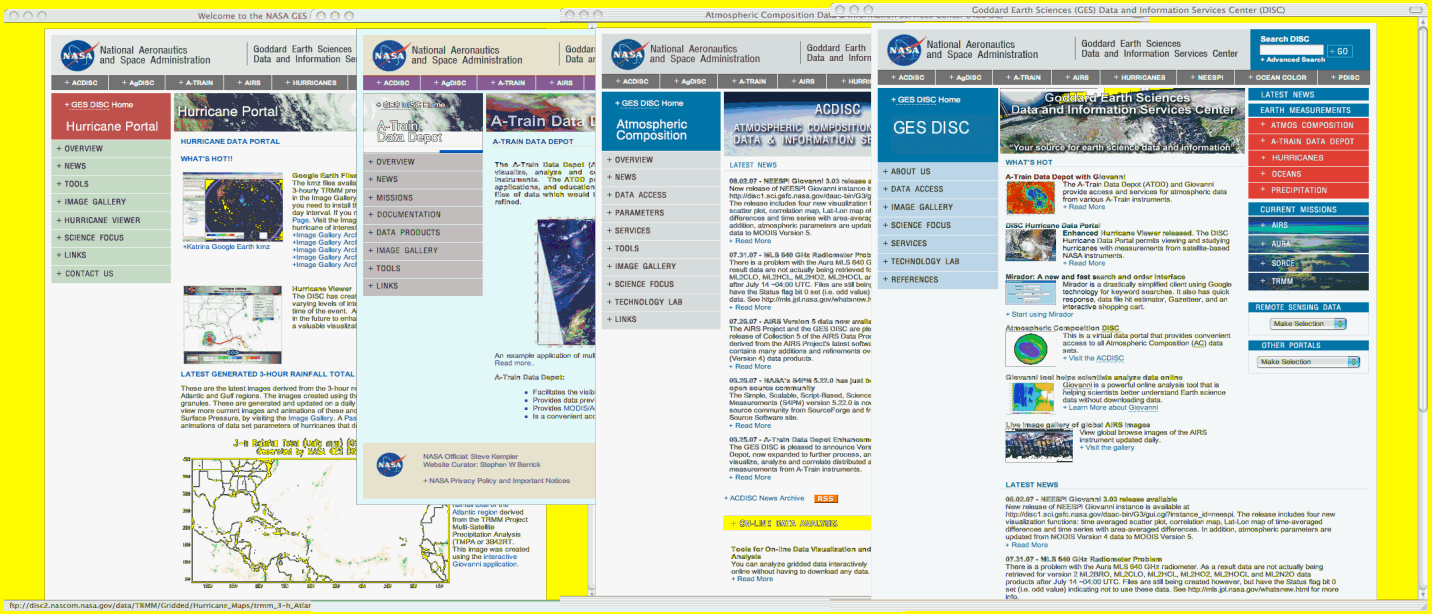
http://s4pm.sci.gsfc.nasa.gov



Data & Information Web Portals

Community and project based portals
Accessible from http://disc.gsfc.nasa.gov
Tailored to the users being served

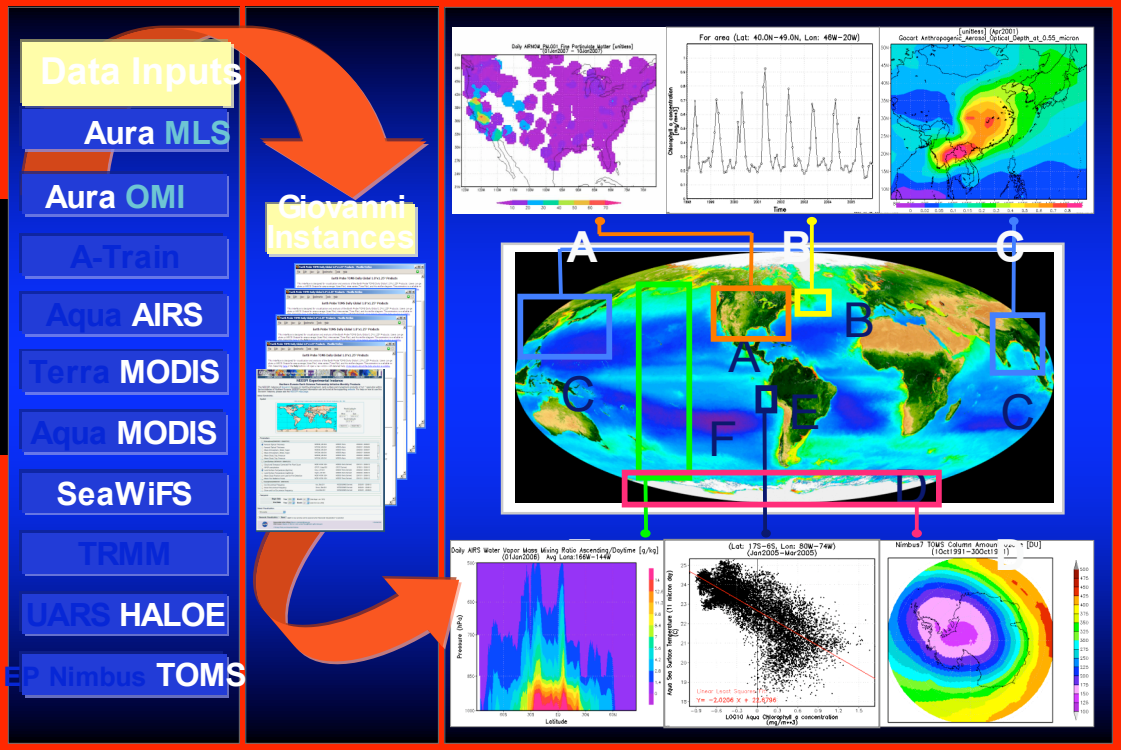
- Multi-mission science research
- Discipline specific portals
- Remote data access



Mirador Data Search

- http://mirador.gsfc.nasa.gov
- Based on Google
- Fast, easy to use
- Gazetteers for places and events
- Can support portals
- Also available as Web Service

Giovanni: Data Visualization and Analysis



- Data from multiple sensors
- Single- and multi-parameter statistics
- Multiple output formats and protocols
  - JPEG and PNG
  - WMS
  - HDF
  - netCDF
  - ASCII
  - KML for Google Earth
- Multiple input formats & protocols
  - HDF4 and HDF5
  - OPeNDAP
  - WCS
  - Web Services (including SciFlo)
  - Standard FTP
- Data lineage support (in development)

http://giovanni.gsfc.nasa.gov

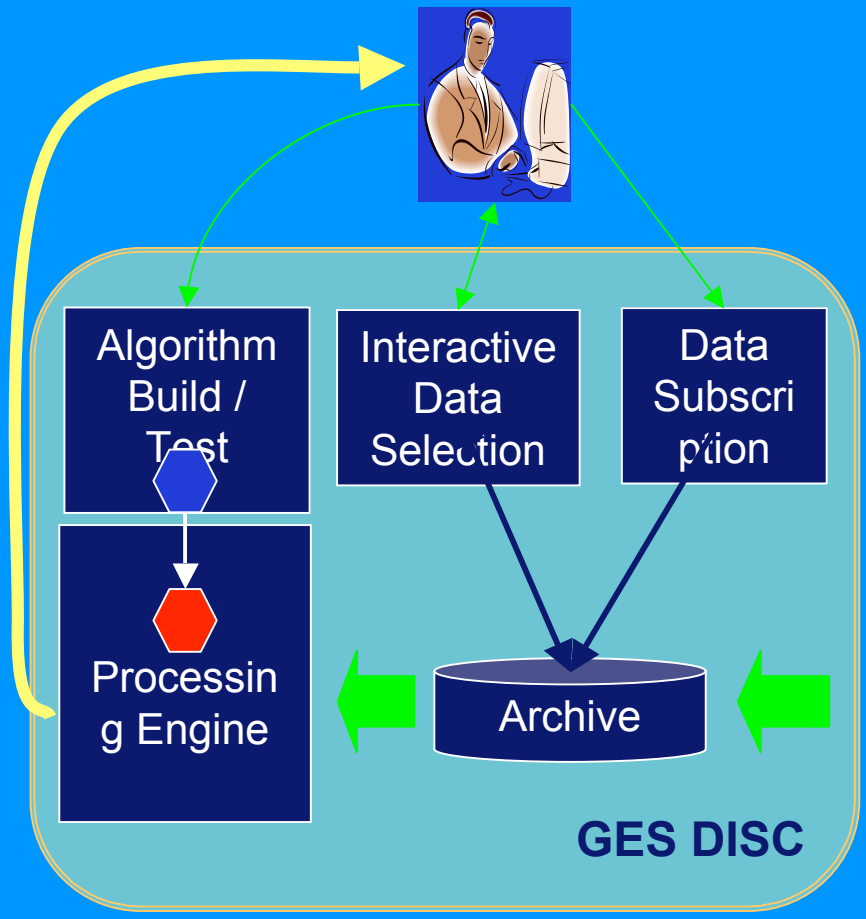
Data Mining

Data mining services available in S4PM

- Users submit and execute data mining algorithms
- Simple Web interface
- Subscriptions process new data as they arrive
- Mining results are made available to user via FTP

Coming Soon: Mining Web Services

- ADaM mining algorithms from Univ. Alabama-Huntsville
- Invoked via Web Services interface

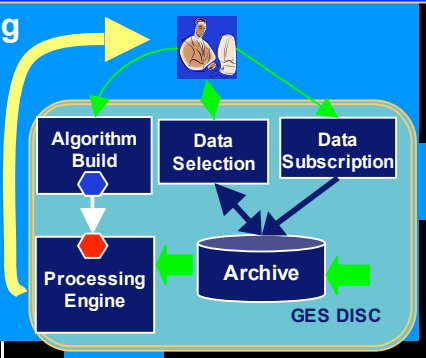


Core Technology Components

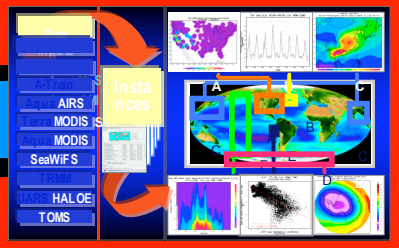
User Community

Science Investigators/Partners

Data Mining



Giovanni: Data Visualization and Analysis



http://giovanni.gsfc.nasa.gov

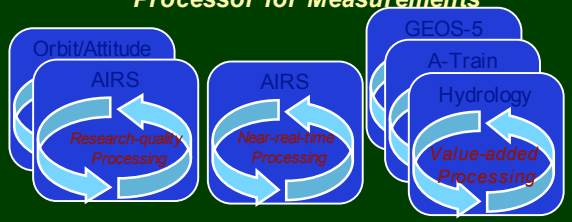
- Data from multiple sensors
- Single- and multi-parameter statistics
- Multiple output formats & protocols
- Data lineage support (in development)

Data & Information Web Portals

- Community and project based portals
- Tailored to the users being served

Data Processing with *S4PM*

Simple, Scalable, Script-based Science Processor for Measurements

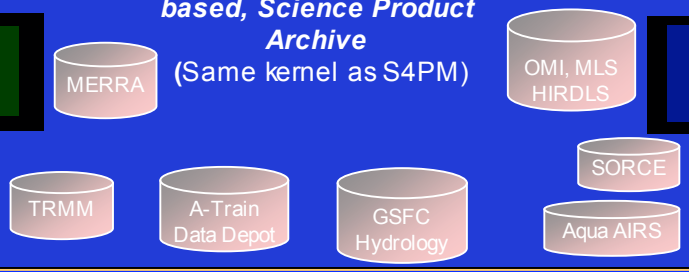


http://s4pm.sci.gsfc.nasa.gov

Data Archive and Distribution with *S4PA*

Simple, Scalable, Script-based, Science Product Archive

(Same kernel as S4PM)



Mirador Data Search

Science Teams

User Community

Atmospheric Composition

Atmospheric Composition Community

Aerosol Researchers, DataFed

Archive
S4PA

OMI, HIRDLS, MLS Teams

MERRA

Climate Modeling Community

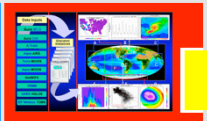
Archive
S4PA

GMAO Science Team

TRMM/Hurricanes

Precipitation Community

Science Investigators



Archive
S4PA

TRMM Science Team

AIRS

Atmospheric Dynamics Community

SPoRT, UMBC et al.

L0 → L3

NRT

Processing
S4PM

Archive
S4PA

AIRS Science Team

A-Train Data Depot

A-Train Constellation Community

Science Investigators

MODIS & OMI
co-location w/
CloudSat &
CALIPSO
S4PM

Archive
S4PA

CloudSat and CALIPSO Science Teams

Hydrology

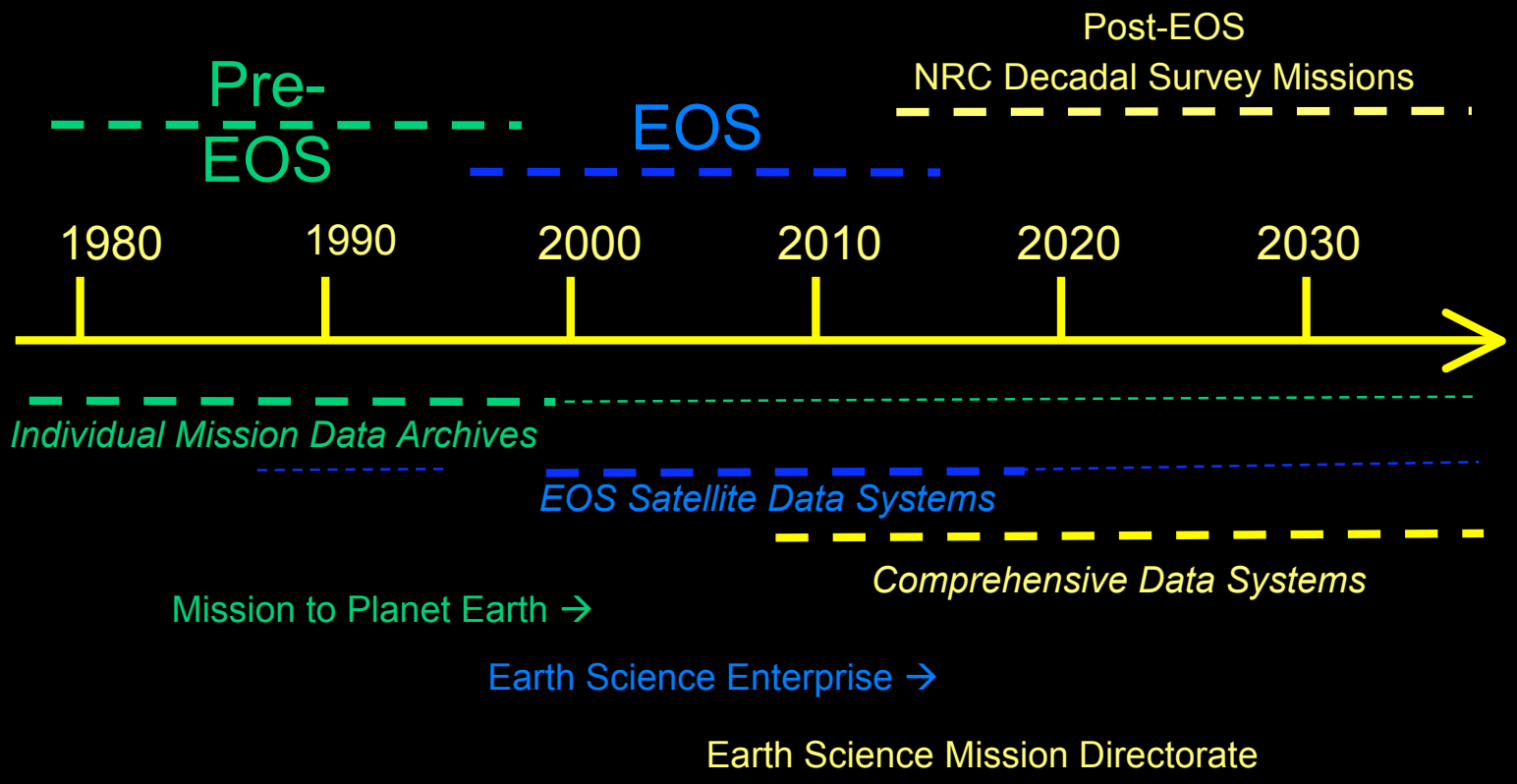
Hydrology Community

GSFC Hydrologic Sciences Branch

Processing
S4PM

Archive
S4PA

Data System Evolution Overview



EOSDIS Evolution at the GES DAAC: 2009 – 2013

Evolved System Characteristics - S4PA based	Evolved System Characteristics - S4PA based initially
Discipline-specific interfaces in addition to generalized interface(s)	Discipline-specific, multi-mission, services oriented interfaces in addition to generalized interface(s)
Disk archive - Some products processed on demand (virtual products) - Download data automatically	Disk archive - Access provided to all products archived locally and remotely - Download data automatically upon choosing
Tools to find, explore, and analyze data	Web-based tools to provide for comprehensive discovery, access, visualization and analysis of coherently related Earth science datasets (satellite, ground-based and model output) to enable interdisciplinary Earth science research
Distribute lower volume tailored products	- Distribute lower volume tailored products - Broker hard to get data
System changes implemented quickly according to priority	System changes implemented quickly according to priority within given budget cap
Steward data	Steward data

Challenges

- How can GES DISC tools be reused for new missions to reduce mission costs, while providing the flexible system needed to facilitate science?
- How can GES DISC expertise be utilized in the formulation of new missions?
- Are the GES DISC work directions consistent with the objectives of the NASA Earth science program and its research communities?
- Are the GES DISC work directions consistent with the objectives of the NASA Earth science program and its research communities?
- How much effort should the GES DISC expend on Earth science applications and application research?
- Integrating information services as a commodity item in the science community.

BACKUP SLIDES

Definitions

- **Data Center** - Ingests, processes, archives, distributes, and manages data acquired from remote sensing instruments. Data "has no significance beyond its existence" *
- **Information Center** - Ingests, processes, archives, distributes, and manages value-added data products, and value-added data processes. "Data that has been given meaning by way of relational connection" *
- **Knowledge Center** (not yet) - "Knowledge is the appropriate collection of information, such that its intent is to be useful" *
- **Data and Information Services Center** - Provides tools that further the use and usefulness of data and information

Each further enables researchers to do their work: Science

* G.Bellinger, et al, "Data, Information, Knowledge, and Wisdom, 2004

GSFC Organizations
(and those most relevant to the GES DISC)

- Code 100 - Office of the Director
- Code 200 - Management Operations Directorate
- Code 300 - Office of Systems Safety and Missions Assurance
- Code 400 - Flight Projects Directorate
  - Code 407 - Earth Science Technology Office
  - Code 420 - Earth Sciences Projects Division
  - Code 423 - Earth Science Data and Information Systems (ESDIS) Project Office
- Code 500 - Applied Engineering and Technology Directorate
  - Code 580 - Information Systems Center
- Code 600 - Sciences and Exploration Directorate
  - Code 610 - Earth Sciences Division
- Code 700 - Information Technology and Communications Directorate
- Code 800 - Suborbital and Special Orbital Projects Directorate

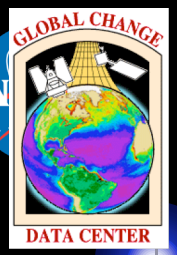
GSFC Organizations

(and those most relevant to the GES DAAC)

Code 610 - Earth Sciences Directorate

- Code 610.1 - Global Modeling and Assimilation Office
- Code 610.2 - Global Change Data Center
- Code 610.3 - Software Integration and Visualization Office
- Code 611 - Goddard Institute of Space Sciences
- Code 613 - Laboratory for Atmospheres
- Code 614 - Hydrospheric and Biospheric Sciences Laboratory

Code 610.2 Organization Chart



Global Change Data Center
S. Wharton, Chief

D. Tabb – Resources Mgr.

Approved ORIGINAL SIGNED BY Date Feb. 18, 1999
Stephen Wharton
Chief
Global Change Data Center
Updated: 4/1/09

GES DISC
S. Kempler, Manager

G. Alcott – Operations Manager
G. Leptoukh – Data Manager
C. Lynnes – Chief System Engineer
P. Smith – Data Support Team Lead
L. Pham- Engineer- Procurement
B. Vollmer - Science Integration
G. Wade – Security Office

GCMD
L. Olsen, Manager

TSDIS
E. Stocker, Manager

C. Cosner – Facility & System Manager
M. McCumber – Dep. TSDIS & Sci. Data Manager

GMAO
G. Kim, Manager

Earth Science Data Operations
R. Kiang, Group Leader

Code 614

SeaWiFS Data System*
G. Feldman, Manager

N. Kuring – Scientific Programmer

* Subject to quality policy of the project organization

Multi-sensor project support

Example: A-Train Data Depot

Currently provides data from Cloudsat, CALIPSO, MODIS (subsets), OMI (subsets), AIRS (subsets)

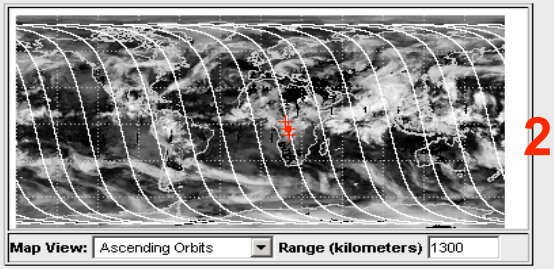
Planned: MLS, HIRDLS, TES, POLDER, CERES, atmospheric model data

CloudSat, and coregistered MODIS/Aqua, AIRS/Aqua, CALIPSO lidar, and OMI/Aura Atmospheric Measurements

Plots of vertical profiles of clouds, temperature, humidty, cloud and aerosol classification; Horizontal swaths of cloud characteristics and total column aerosols, collocated with CloudSat track; Line over-plots of cloud pressures.

Select Constraints:

Spatial



Map View: Ascending Orbits | Range (kilometers): 1300

2. Select Location

Temporal

Orbit Date | Year: 2007 | Month: Sep | Day: 6 | Update Map | *(Range: 02 Jun 2006 - 11 Sep 2007)*

1. Pick Date

Parameters

Curtains

Temperature(2002/08/30 - 2007/09/19)			
☐ Atmospheric Temperature Profile	AIRX2RET.003	AIRS Aqua	2002/08/30 - 2007/09/01
☐ Atmospheric Temperature Profile	AIRX2RET.005	AIRS Aqua	2002/08/31 - 2007/09/16
☐ Atmospheric Temperature Profile (Kelvins)	MAC07S0.002	MODIS Aqua	2006/06/02 - 2007/09/19
Water Vapor(2002/08/30 - 2007/09/19)			
☐ H2O (Dew_Point_Temperature_Profile in Kelvins)	MAC07S0.002	MODIS Aqua	2006/06/02 - 2007/09/19
☐ H2O Saturation Mass Mixing Ratio (gm/kg dry air)	AIRX2RET.005	AIRS Aqua	2002/08/31 - 2007/09/16
☐ H2O Saturation Mass Mixing Ratio (gm/kg dry air)	AIRX2RET.003	AIRS Aqua	2002/08/30 - 2007/09/01
☐ H2O Vapor Mass Mixing Ratio (gm/kg dry air)	AIRX2RET.003	AIRS Aqua	2002/08/30 - 2007/09/01
☐ H2O Vapor Mass Mixing Ratio (gm/kg dry air)	AIRX2RET.005	AIRS Aqua	2002/08/31 - 2007/09/16
Clouds(2006/06/02 - 2007/09/17)			
☑ Cloud/Aerosol Classification (Vertical Feature Mask)	VFM.001	Calipso - Lidar	2006/06/13 - 2007/09/17
☐ ReceivedEchoPowers	1B_CPR.008	CloudSat	2006/06/02 - 2007/09/12
☑ Reflectivity dBZ	1B_CPR.008	CloudSat	2006/06/02 - 2007/09/12
☐ RO Ice Water Content	2B_CWC_RO.007	CloudSat	2007/01/07 - 2007/01/08
☐ RO Liquid Water Content	2B_CWC_RO.007	CloudSat	2007/01/07 - 2007/01/08

Strips

Surface(2002/08/30 - 2007/09/20)			
☑ Cloud Top Pressure in hPa (Horizontal Strip)	MAC06S1.002	MODIS Aqua	2006/06/02 - 2007/09/16
☑ Aerosol Optical Depth 550nm	MAC04S1.002	MODIS Aqua	2006/06/02 - 2007/09/16
☑ Effective Cloud Pressure for O3 (Raman Ring)	OMCLRRS0.001	OMI Aura	2007/08/11 - 2007/08/12
☑ Effective Cloud Pressure (O2-O2)	OMCLO2S0.001	OMI Aura	2006/06/02 - 2007/09/18
☑ Final Aerosol Absorption Optical Depth	OMCLAERUVS0.002	OMI Aura	2006/06/02 - 2007/09/24
☑ UV Aerosol Index	OMCLTO3S0.002	OMI Aura	2006/06/02 - 2007/09/22

3. Choose Parameters

Select Visualization:

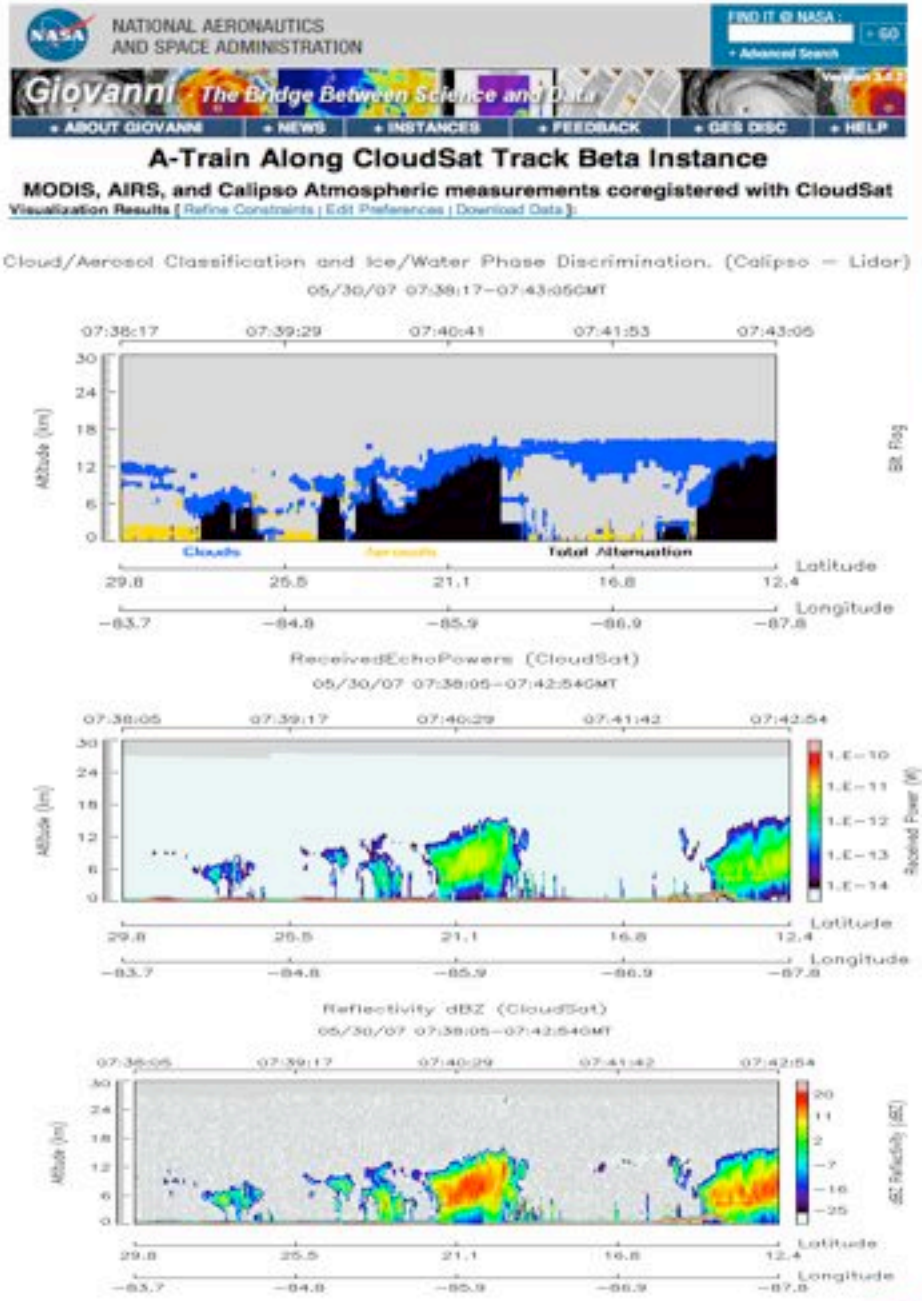
Subset Parameters Along Orbit Track - H2P

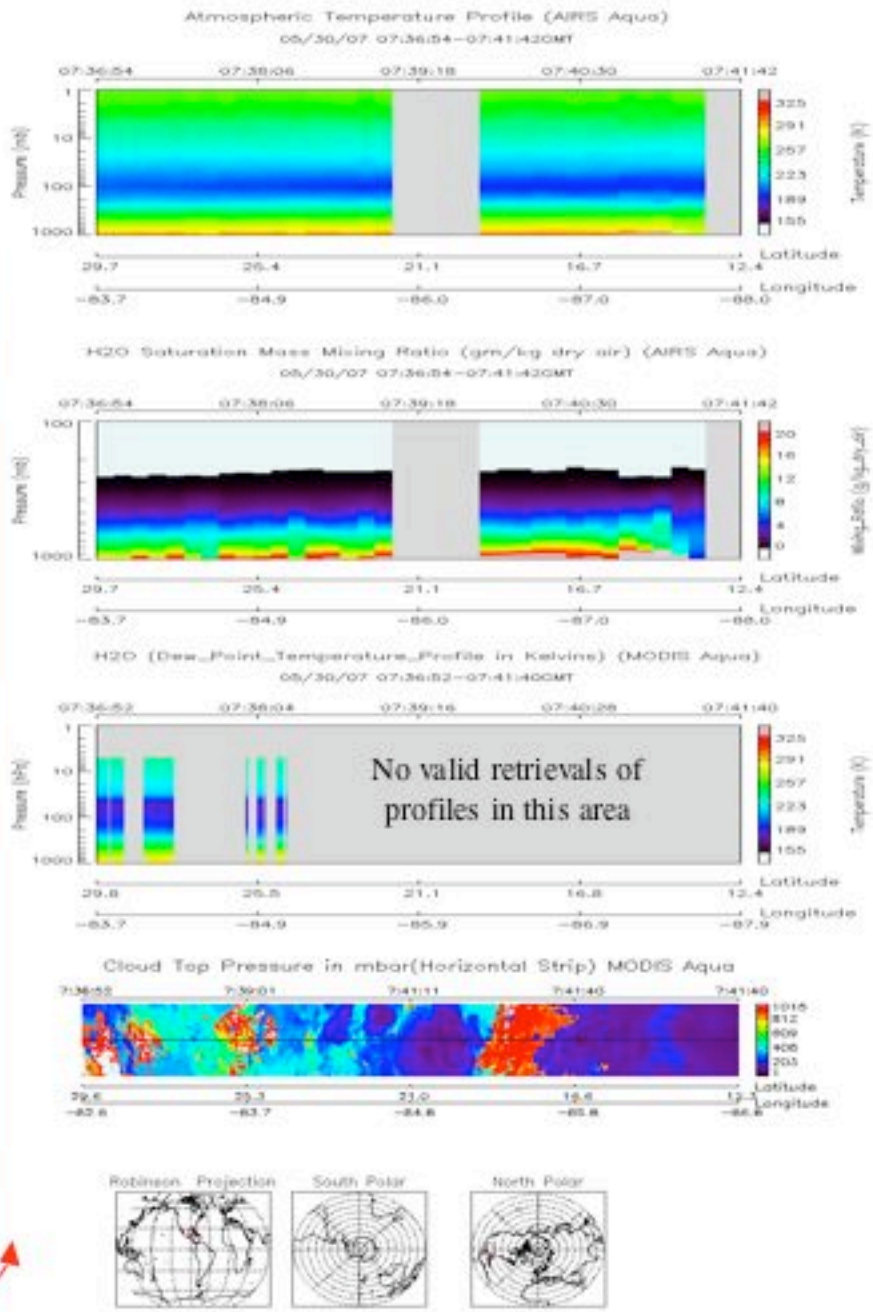
4. Generate Visualization

Services Help

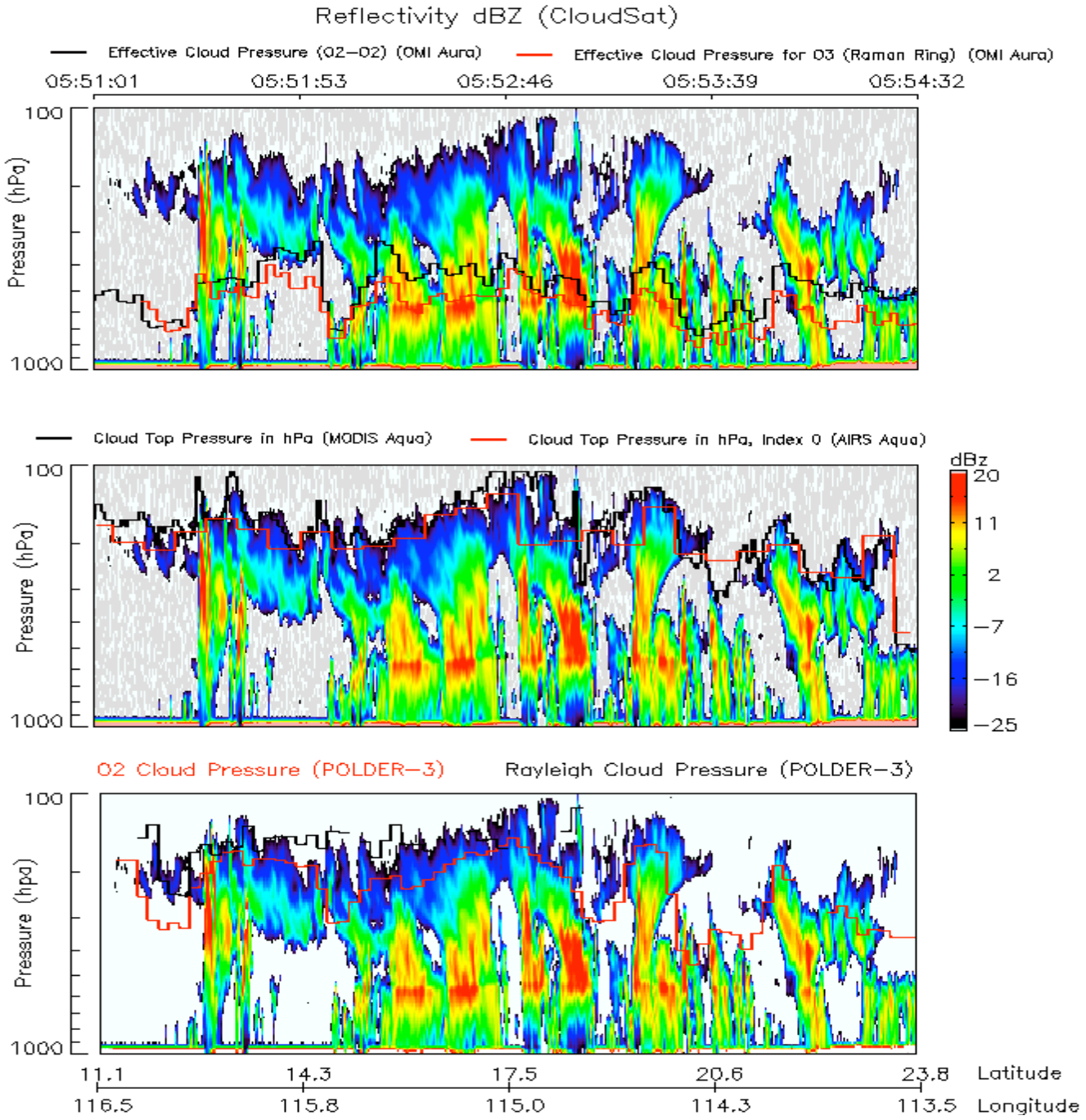
Generate Visualization | Reset | ***Alert:*** *A new window will be opened when "Generate Visualization" is selected.*

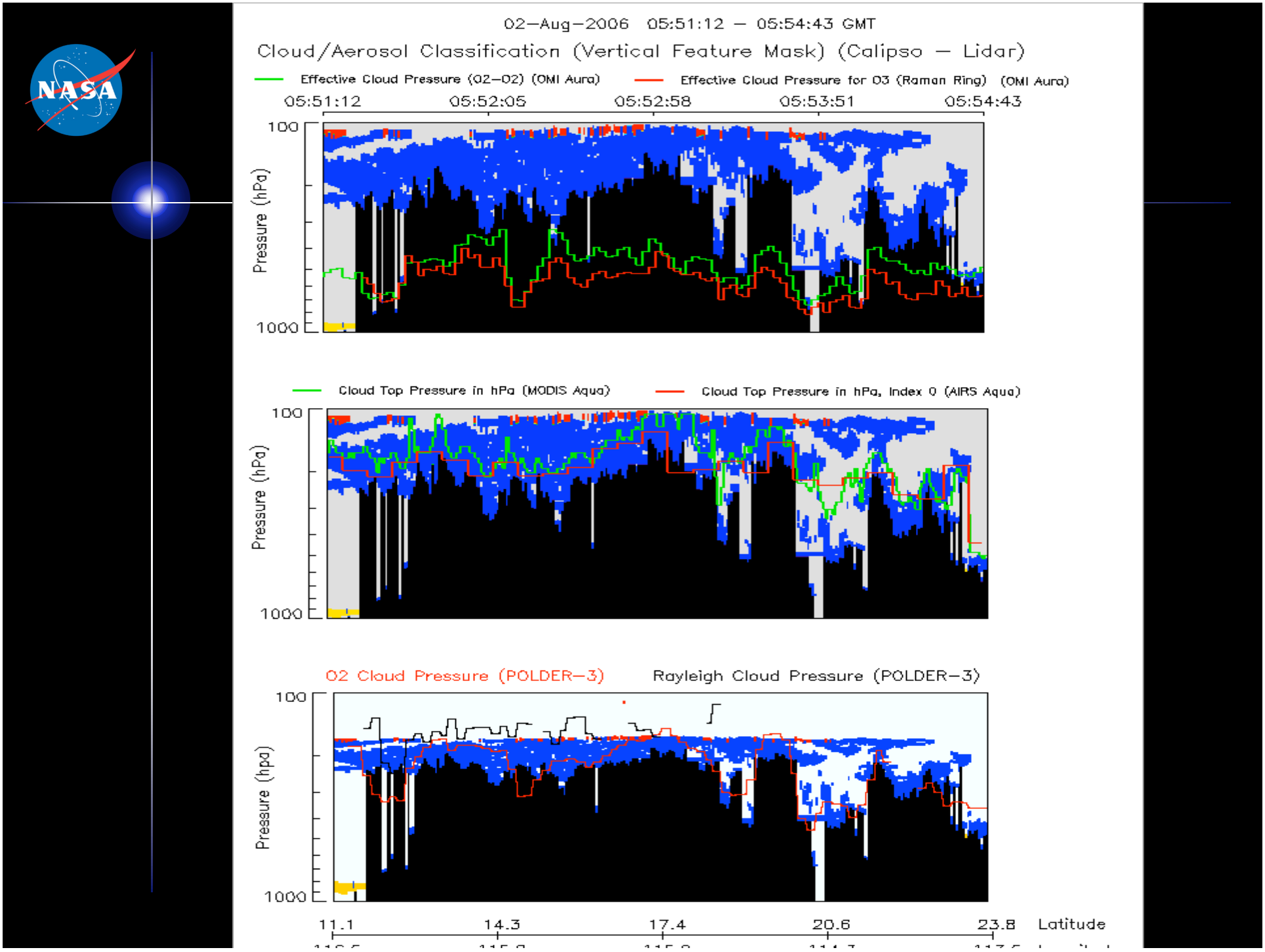
Tropical Storm Barry, May 30, 2007, forming in the Gulf of Mexico



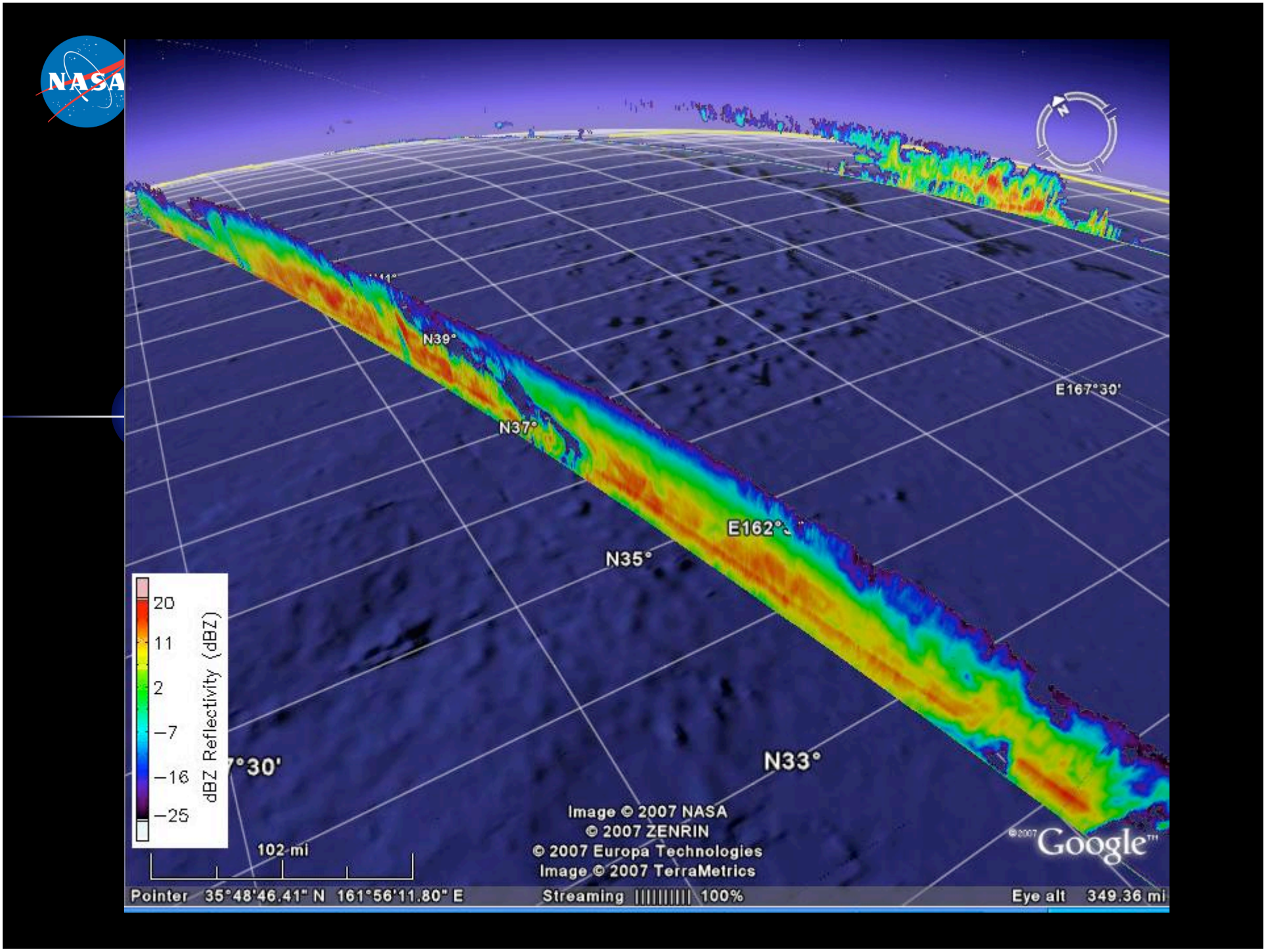


02–Aug–2006 05:51:01 – 05:54:32 GMT



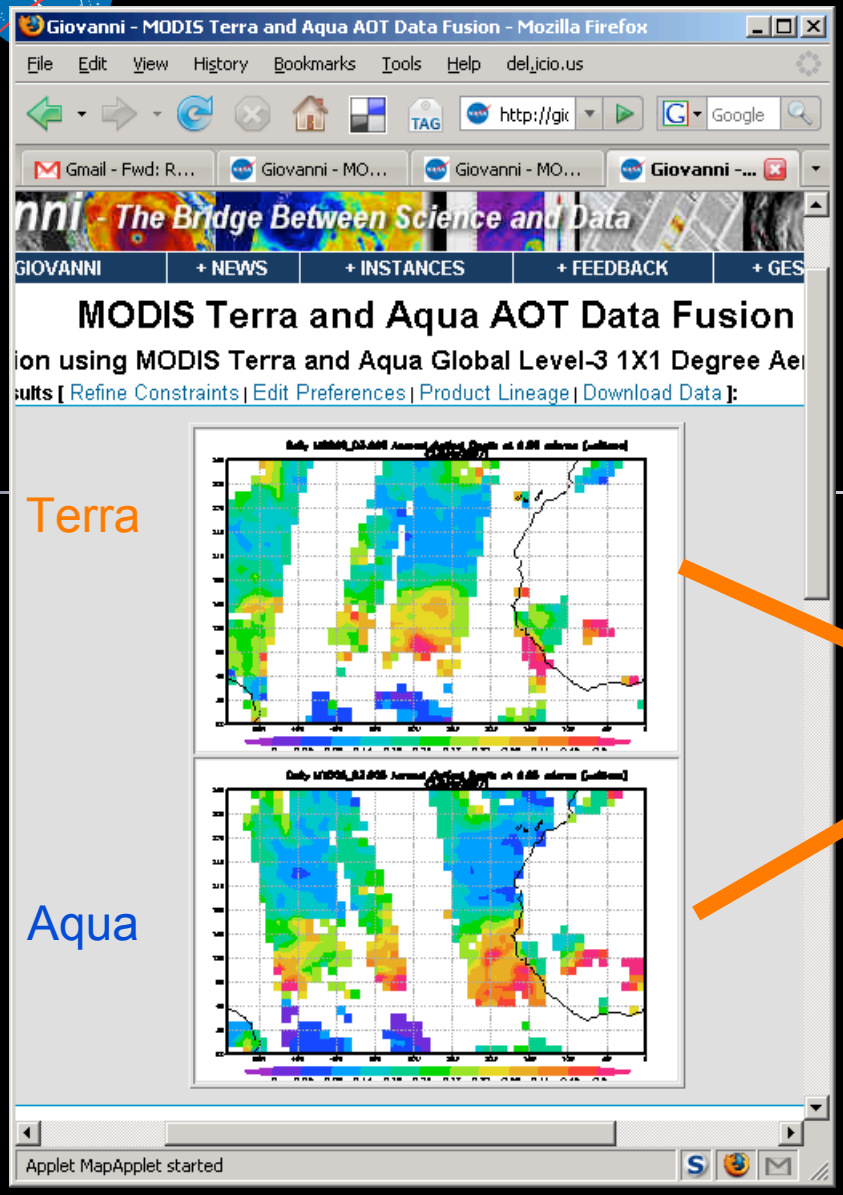

02-Aug-2006 05:51:12 - 05:54:43 GMT
Cloud/Aerosol Classification (Vertical Feature Mask) (Calipso - Lidar)
Effective Cloud Pressure (O2-O2) (OMI Aura)
Effective Cloud Pressure for O3 (Raman Ring) (OMI Aura)
05:51:12
05:52:05
05:52:58
05:53:51
05:54:43
100
1000
Pressure (hPa)
Cloud Top Pressure in hPa (MODIS Aqua)
Cloud Top Pressure in hPa, Index 0 (AIRS Aqua)
O2 Cloud Pressure (POLDER-3)
Rayleigh Cloud Pressure (POLDER-3)
Pressure (hpa)
11.1
14.3
17.4
20.6
23.8
Latitude

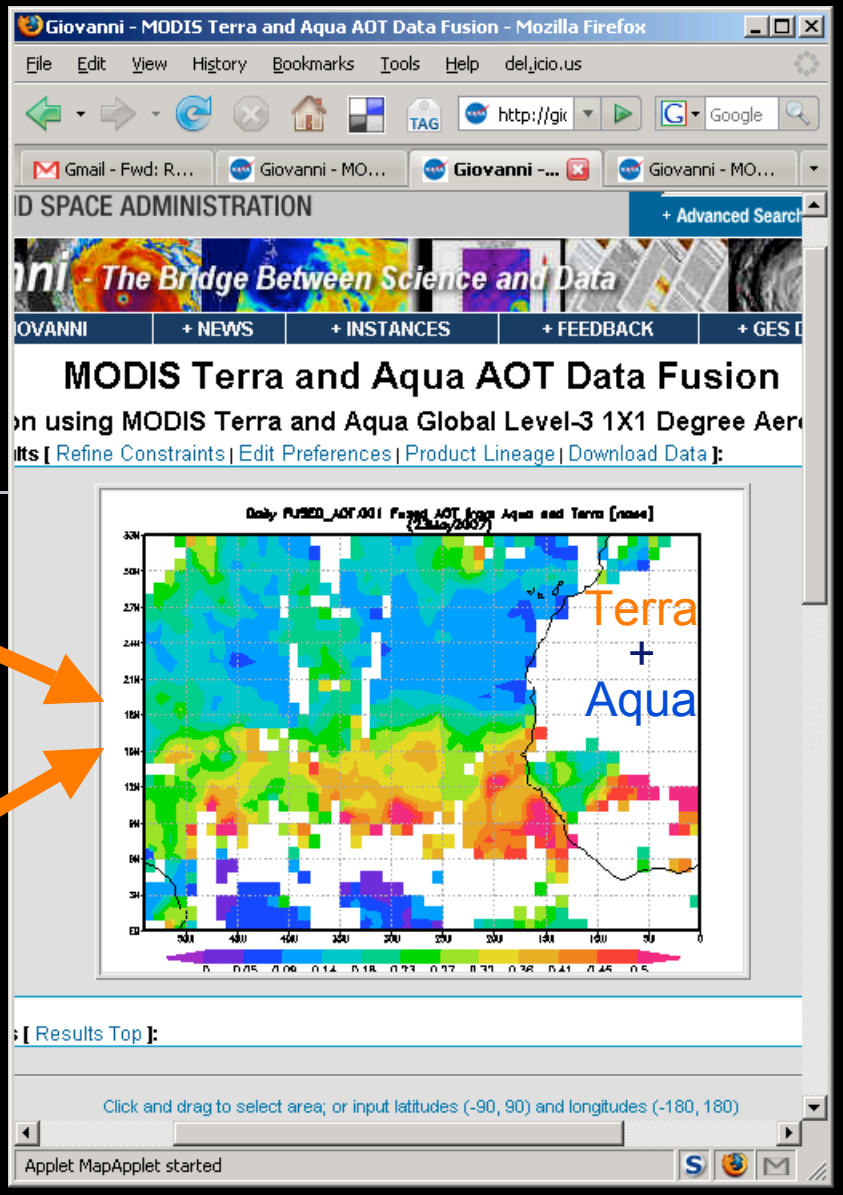
Google Earth

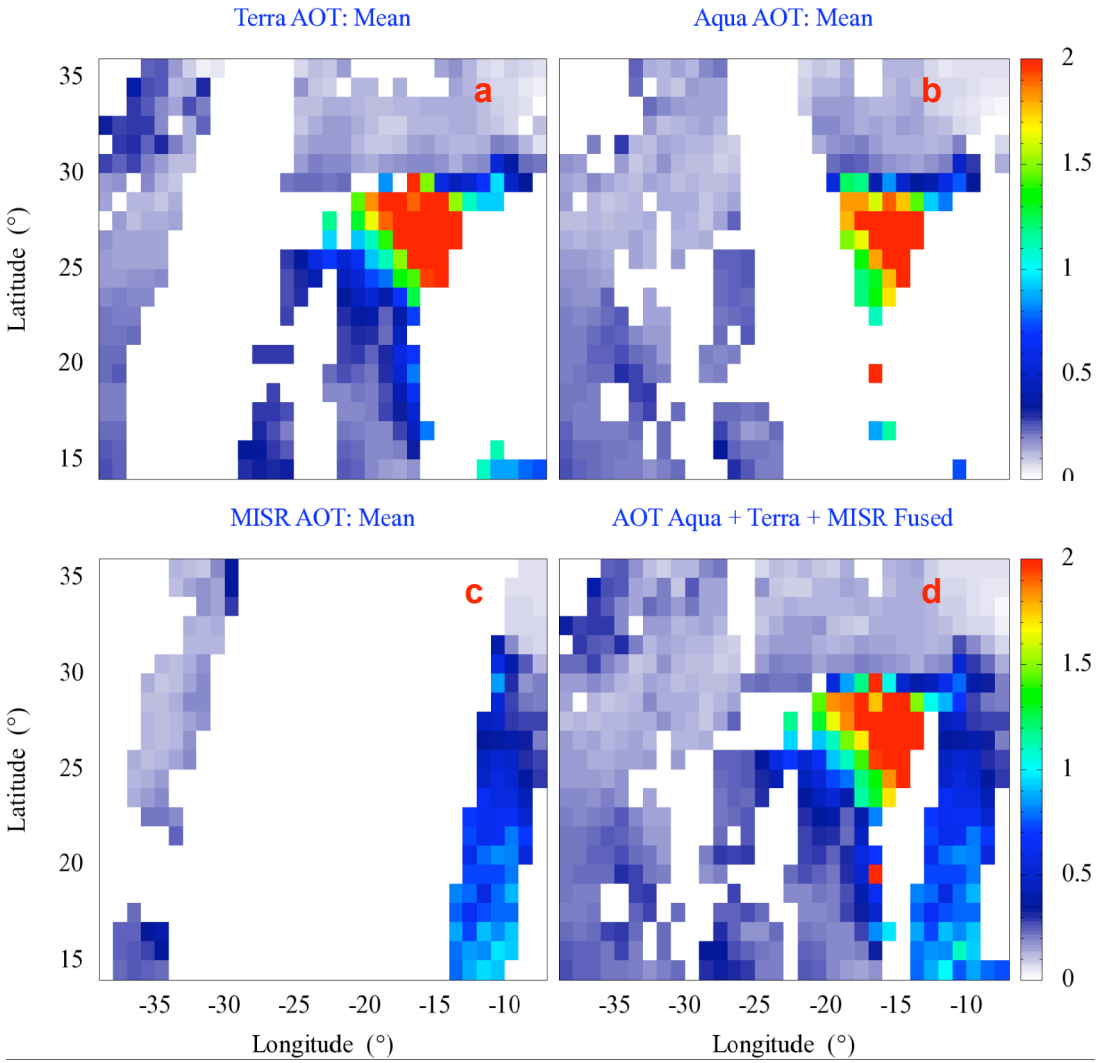

N39°
N37°
E167°30'
E162°
N35°
N33°
20
11
2
−7
−16
−25
dBZ Reflectivity (dBZ)
102 mi
Image © 2007 NASA
© 2007 ZENRIN
© 2007 Europa Technologies
Image © 2007 TerraMetrics
Google
Pointer 35°48'46.41" N 161°56'11.80" E
Streaming 100%
Eye alt 349.36 mi

Data Merging

Data Fusion in Giovanni (prototype)





Dust event, May 23, 2007

Next step: Adding MISR